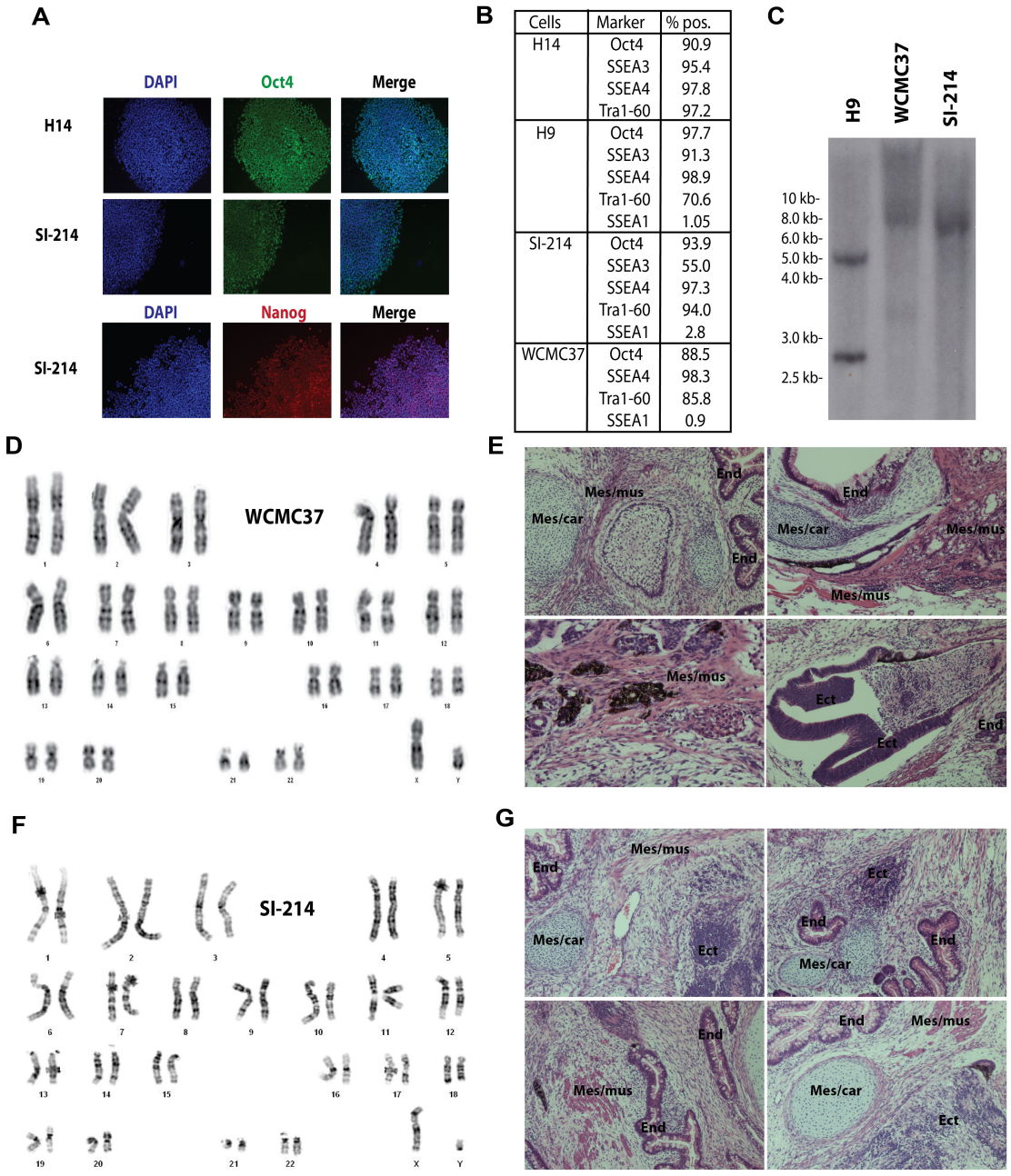
Supplemental Information

The DNA replication program is altered at the *FMR1* locus in fragile X embryonic stem cells

Jeannine Gerhardt, Mark J. Tomishima, Nikica Zaninovic, Dilek Colak, Zi Yan, Qiansheng Zhan, Zev Rosenwaks, Samie R. Jaffrey and Carl L. Schildkraut

Supplemental Information Inventory



Cells	Marker	% pos.
H14	Oct4	90.9
	SSEA3	95.4
	SSEA4	97.8
	Tra1-60	97.2
H9	Oct4	97.7
	SSEA3	91.3
	SSEA4	98.9
	Tra1-60	70.6
	SSEA1	1.05
SI-214	Oct4	93.9
	SSEA3	55.0
	SSEA4	97.3
	Tra1-60	94.0
	SSEA1	2.8
WCMC37	Oct4	88.5
	SSEA4	98.3
	Tra1-60	85.8
	SSEA1	0.9

Figure S1. Characterization of the human FXS hESC examined in this study, Related to Figure 1

(A, B) Undifferentiated hESC grow in colonies and express the surface markers Tra-1-60, SSEA3 and SSEA4. In addition, hESC are also characterized by the expression of the pluripotency markers Oct4 and Nanog. Using immunohistochemistry (A) and flow cytometry (B) we confirmed that the FXS hESC, SI-214 (Verlinsky et al., 2005) and WCMC37 and non-affected hESC lines H14 and H9 (Thomson et al., 1998) express these pluripotency markers. The markers Tra-1-60, SSEA3 and 4 were present in a high

percent in these cells. Only a very low percent of FXS hESC were positive for SSEA-1, a differentiation marker.

(C) To determine repeat instability we expanded FXS hESC from a single colony and analyzed the cells by Southern blot. DNA was digested with *EcoRI* and *EagI*.

(D-F) FXS hESC lines WCMC37 (D) and SI-214 (F) showed a normal XY karyotype.

(E-G); Teratoma formation: Mice were injected with FXS hESC WCMC37 (E) or SI-214 (G) and after several months the teratomas were histologically analyzed. Both FXS hESC formed teratomas *in vivo*, which contained tissue from ectoderm (Ect), endoderm (End) and mesoderm (Mes); either cartilage (Car) or muscle (Mus).

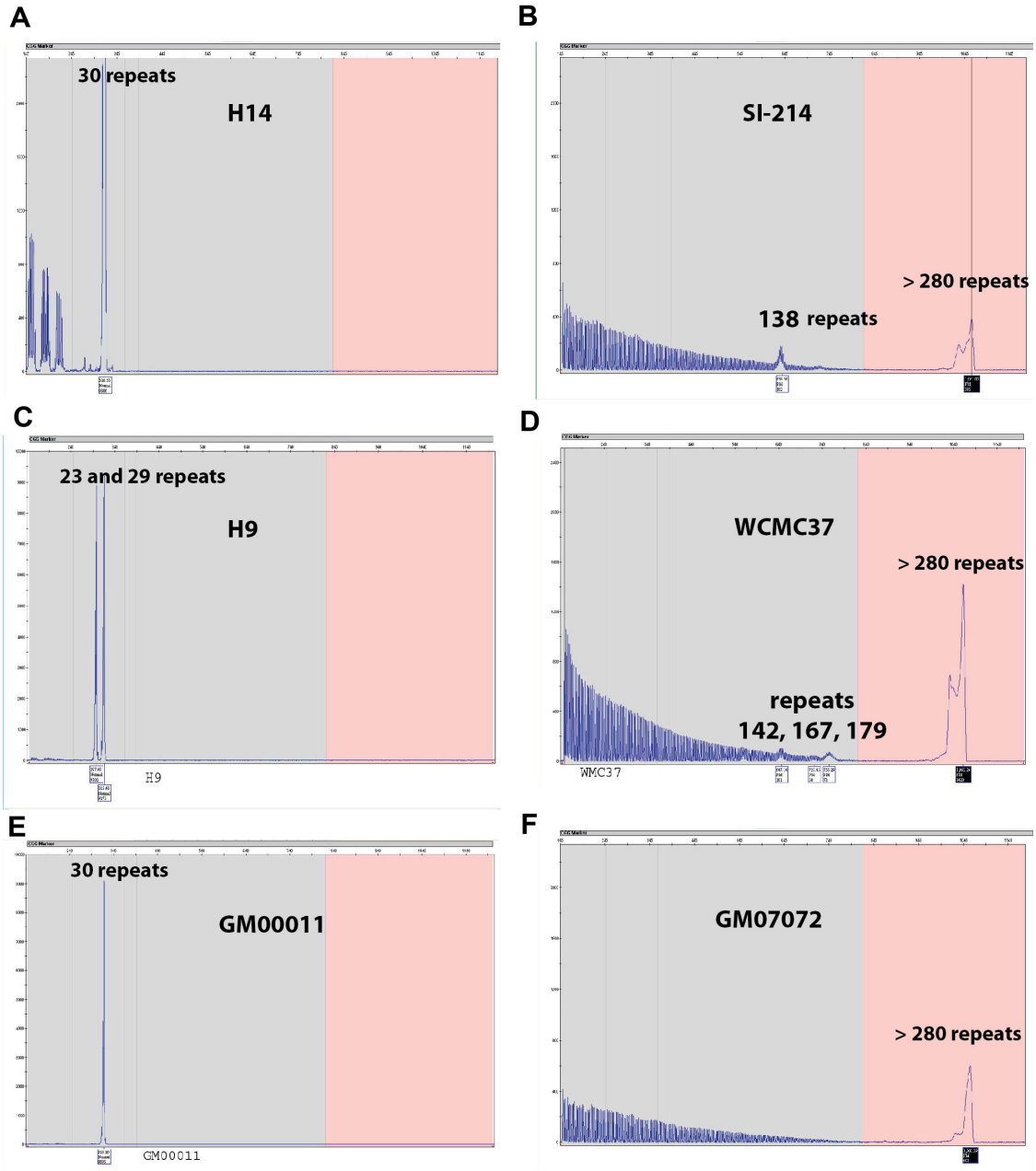


Figure S2. Mosaicism of the repeats in FXS hESC was determined by analysis of the CGG repeat length using the AmplideX™ PCR assay (Asuragen) and capillary electrophoresis, Related to Figure 1

Genomic DNA was isolated from human ESC H14 (A), H9 (C), SI-214 (B), WCMC37 (D), GM00011 (E) and GM07072 (F). The CGG repeat length was measured by PCR and capillary electrophoresis, which determines the repeat length of normal, pre-mutation and full-mutation alleles with an accuracy up to 280 repeats.

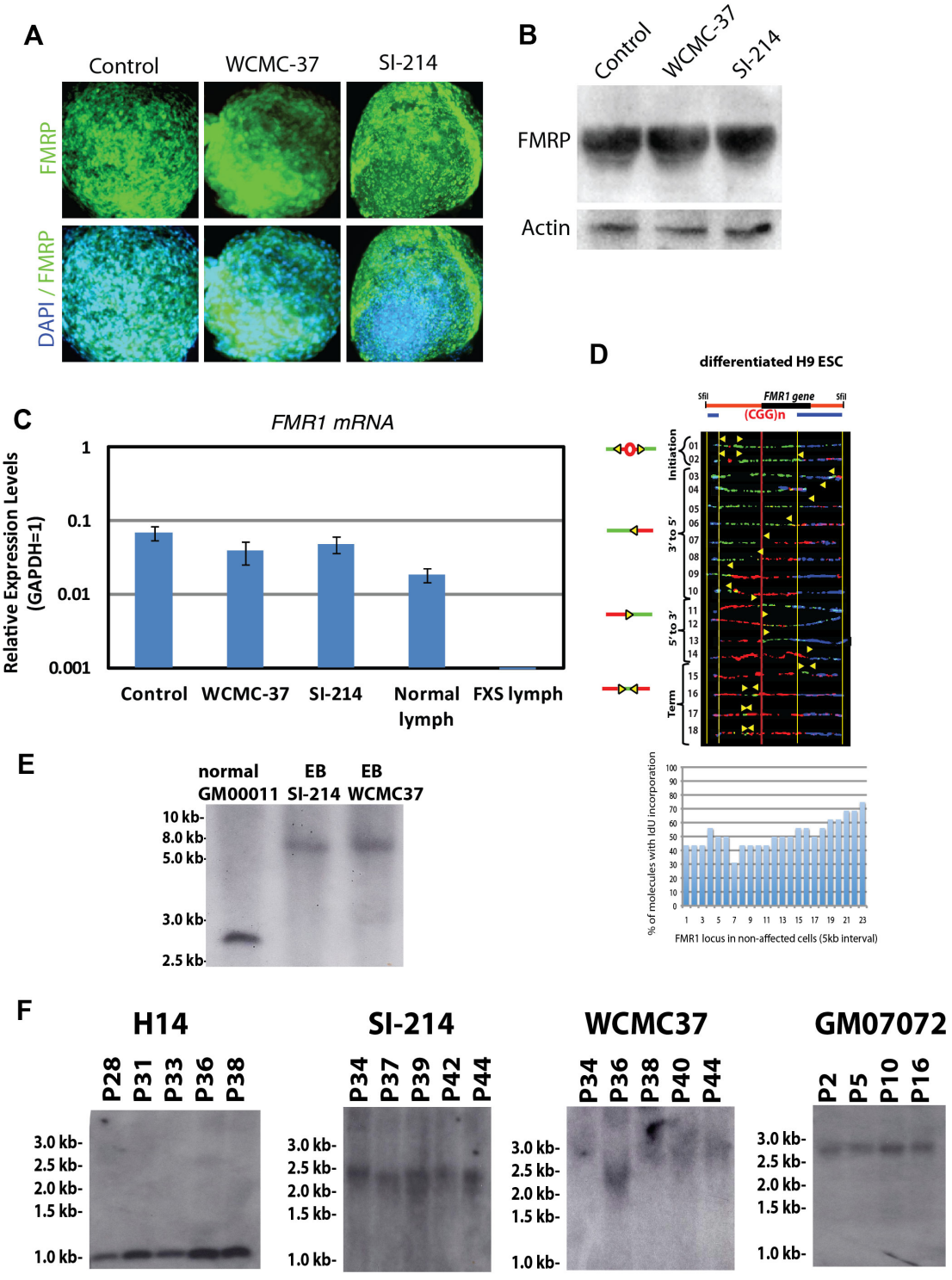


Figure S3. Characterization of the *FMR1* transcription, CGG repeat instability in human FXS hESC and DNA replication profile in differentiated H9 ESC, Related to Figure 2

(A-C) Undifferentiated hESC express FMRP, which we determined by immunostaining (A) and Western blot (B) with FMRP antibody. Further quantification of the *FMR1* mRNA level showed transcription of the *FMR1* gene in FXS hESC, but not in differentiated FXS cells. Data are represented as mean +/- SEM, n=5 biological replicates.

(D) Top; Map of the *FMR1* gene locus with *SfiI* restriction enzyme cutting sites and FISH probes indicated (blue). A vertical red bar marks the CGG repeats. Bottom; Photomicrographs of labeled DNA molecules from non-affected differentiated H9 cells are arranged by replication fork progression in the 5'→3' and the 3'→5' directions, replication initiation, and termination sites. Yellow arrows mark the location of the replication forks in each individual DNA molecule. The percentage of molecules with IdU incorporation (first pulse) is calculated from the DNA molecules shown above.

(E) To determine repeat length in differentiated FXS SI-214 EB and WCMC37 EB, genomic DNA was digested with *EcoRI* and the methylation-sensitive enzyme *EagI* and analyzed by Southern blot.

(F) To determine repeat instability we collected several passages from FXS hESC SI-214, WCMC37 and from the differentiated fetal fibroblast line GM07072. We digested the genomic DNA with *PstI* enzyme and analyzed the DNA by Southern blot.

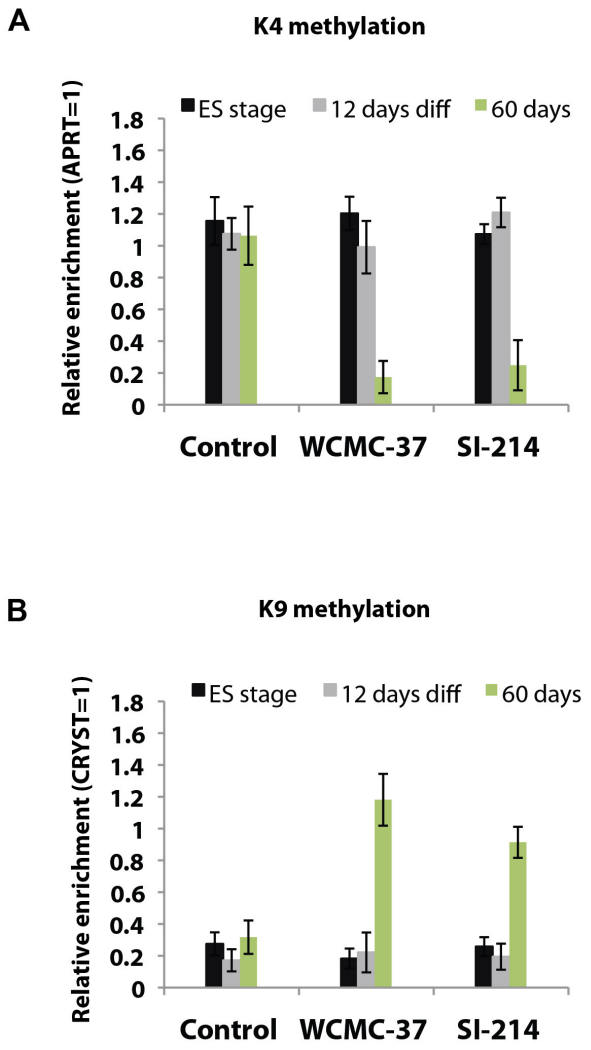


Figure S4. The *FMR1* promoter is enriched for H3K4 methylation and lacks H3K9 methylation in FXS hESC and 12 days after differentiation, Related to Figure 4

(A) To determine the H3K4 methylation status in hESC, 12 and 60 days after differentiation in control and FSX cells SI-214 and WCMC37 we analyzed the enrichment of H3K4 methylation by Chromatin-immunoprecipitation (ChIP). The DNA isolated by ChIP was used in qRT-PCR with primers that are adjacent to the *FMR1* promoter. Data are represented as mean +/- SEM, n=4 biological replicates.

(A) To determine the H3K9 methylation status in hESC, 12 and 60 days after differentiation in control and FSX cells SI-214 and WCMC37 we analyzed the enrichment of H3K9 methylation by ChIP with primers located adjacent to the *FMR1* promoter. Data are represented as mean +/- SEM, n=4 biological replicates.

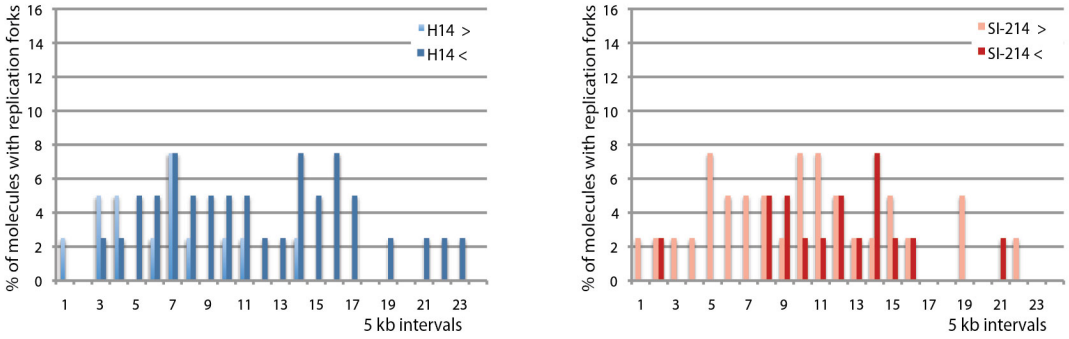


Figure S5. The DNA replication fork does not stall at a specific position in the flanking region upstream of the *FMR1* gene, Related to Figure 5

The percentage of replication forks for each 5 kb DNA interval was determined for the *SfiI* DNA segment (119 kb) upstream of the *FMR1* gene in non-affected hESC H14 (left) and FXS hESC SI-214 (right).

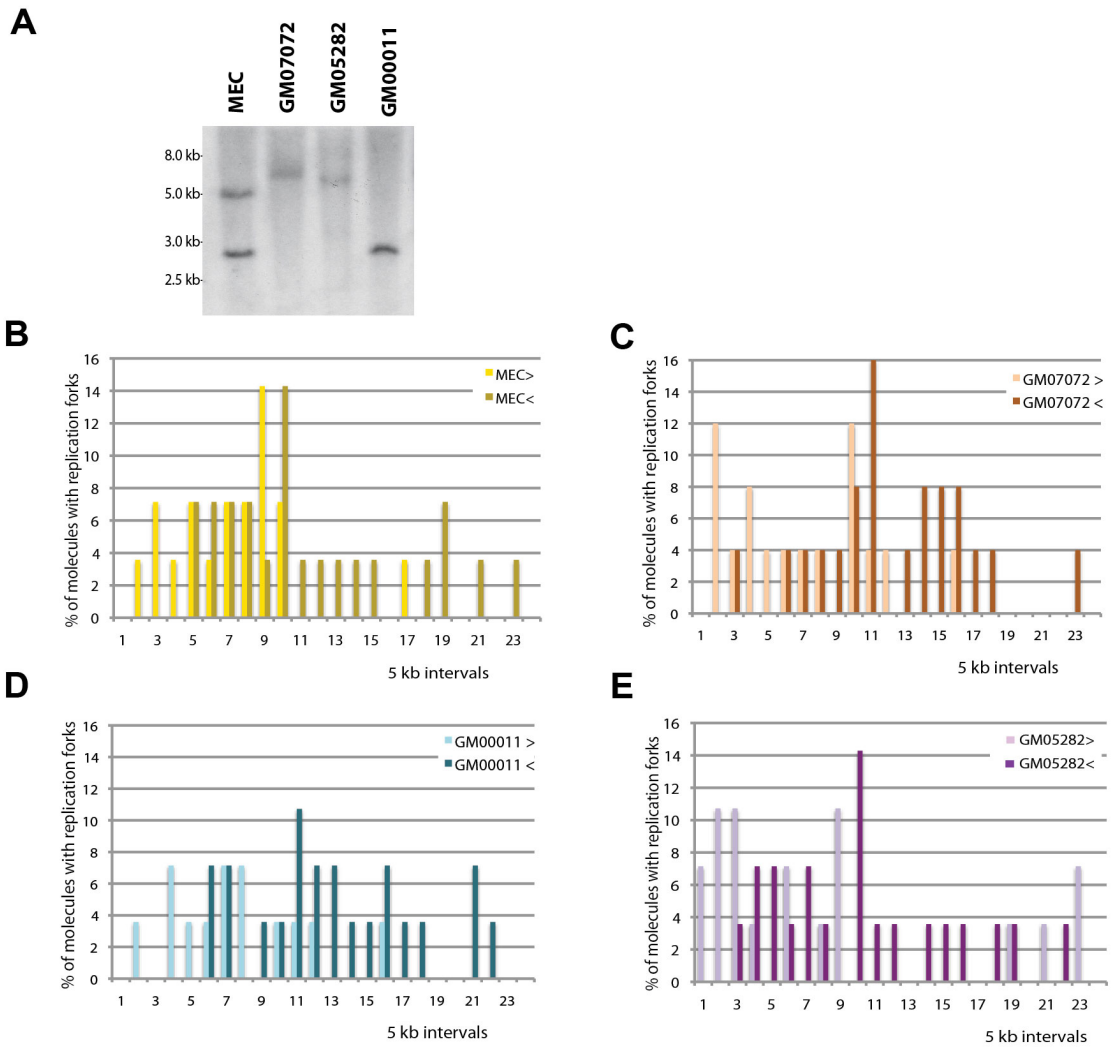


Figure S6. DNA replication forks stall at CGG/CCG repeats in the *FMR1* locus, Related to Figure 6

(A) To determine repeat length, the genomic DNA was digested with *EcoRI* and methylation-sensitive enzyme *EagI* and analyzed by Southern blot.

(B-G) The percentage of molecules with replication forks for each 5 kb DNA interval was counted at the *FMR1* locus in (B) non-affected MEC and (D) fetal fibroblast GM00011 and (C) FXS fetal fibroblast GM07072 and (E) FXS fetal fibroblast GM05282. The replication fork in the 3'-5' direction is indicated by > and in the 5'-3' direction by <.

Cells	DNA fragment	Total molecules	Fork speed from Red (kb/min)	Fork speed from Green (kb/min)	Average Fork speed (kb/min)	% difference between fork speeds
H14	SfiI 120	160	1.28	1.05	1.16	10.0
	downstream		1.17	0.91	1.04	12.2
	upstream		0.92	0.71	0.81	12.8
	SbfI 152	108	0.97	1.02	1.00	2.7
	SfiI 119	141	1.06	0.94	1.00	6.1
	PmeI 102	66	1.19	1.13	1.16	2.4
SI-214	SfiI 120	261	1.73	1.69	1.71	1.0
	downstream		2.30	2.24	2.27	1.2
	upstream		1.03	1.00	1.02	1.1
	SbfI 152	108	1.03	1.20	1.12	7.6
	SfiI 119	169	1.03	1.23	1.13	8.6
	PmeI 102	68	1.19	1.13	1.16	2.4
H9	SfiI 120	263	1.63	1.77	1.70	4.3
	downstream		1.69	1.88	1.78	5.0
	upstream		1.05	1.16	1.10	4.9
	PmeI 102	247	1.34	1.60	1.47	9.0
WCMC37	SfiI 120	179	1.26	1.28	1.27	0.9
	downstream		1.14	1.17	1.15	1.1
	upstream		0.92	0.94	0.93	1.1
	PmeI 102	185	1.56	1.37	1.47	6.5
GM00011	SfiI 120	164	1.40	1.26	1.33	5.2
	downstream		1.33	1.18	1.25	5.8
	upstream		1.62	1.19	1.40	15.4
MEC	SfiI 120	179	1.31	1.32	1.31	0.5
	downstream		1.75	1.77	1.76	0.5
	upstream		1.01	1.02	1.01	0.4
GM07072	SfiI 120	173	1.32	0.97	1.14	15.4
	downstream		2.36	1.66	2.01	17.3
	upstream		0.89	0.63	0.76	17.7
GM05282	SfiI 120	129	1.05	0.95	1.00	5.1
	downstream		1.58	1.39	1.48	6.6
	upstream		0.63	0.56	0.60	5.8
WCM37EB	SfiI 120	90	1.16	1.08	1.12	3.6
	downstream		1.14	1.05	1.09	4.4
	upstream		0.96	0.88	0.92	4.5
SI-214 EB	SfiI 120	63	1.34	1.09	1.22	10.3
	downstream		1.38	1.04	1.21	13.8
	upstream		0.88	0.69	0.79	11.8

Table S1. Calculation of the replication fork speed (Norio and Schildkraut, 2001) at the *FMR1* locus (*SfiI* segment containing the *FMR1* gene, Fig. 1), Related to Figure 4

All DNA molecules obtained by SMARD are the total molecules (first column). The average number of replication forks per segment is used to calculate the average speed per replication fork (kb/min) using the following equation: **Average kb /min = [Length of segment (kb) / (Td (min)] / Average number of replication forks for this segment).** Td; duplication time of the fragment; the time required for the segment to duplicate (min), which is calculated by the following equation: **Td= Tp1 x NRG / (NR + NRG) or Td= Tp1 x NRG / (NG + NRG).** Tp1 = Time for the first or second labeling (4 hours). NR or NG = the number of molecules fully stained in red or green. NRG = the number of molecules fully stained in both red and green in the segment. We determined the replication fork speed in two different ways, which agreed in most cases within less 10%. At least 2,700 molecules were analyzed for these calculations. For the 18 different segments we assessed the % difference from the average, which ranged from 0.5 to 15.4 with only 2 values greater than 10% and a mean of 5.7 (3.7, 7.6) suggesting that we have 95% confidence that the true mean % difference was less than 7.6 %.

Cells		*p*-value
H14/H9	normal/normal hESC	0.706
H14/SI214	normal/ FXS hESC	0.00892
H14/WCMC37	normal/ FXS hESC	0.00529
H9/SI214	normal/ FXS hESC	0.00186
H9/WCMC37	normal/ FXS hESC	0.00103
SI214/WCMC37	FXS/FXS hESC	0.414
MEC/GM00011	normal/ normal diff.	0.654
MEC/GM07072	normal/ FXS diff.	0.420
MEC/GM05282	normal/ FXS diff.	0.561
GM00011/GM07072	normal/ FXS diff.	0.417
GM00011/GM05282	normal/ FXS diff.	0.407
GM07072/GM05282	FXS/FXS diff.	0.486
GM07072/SI214	FXS diff/FXS hESC	0.00126
GM05282/SI214	FXS diff/FXS hESC	0.01349
GM07072/WCMC37	FXS diff/FXS hESC.	0.00763
GM05282/ WCMC37	FXS diff/FXS hESC.	0.00820

Table S2. Comparison of replication fork direction (the percentage of IdU incorporation) between the different cell lines by statistically quantifying the chance that the replication profiles are the same or differ between two cell lines, Related to Figure 2

The *p*-value if < 0.02 shows that there is a significant difference between the two replication fork progressions/directions in the two compared cell lines. A *p*-value of > 0.02 shows, that these replication profiles are very similar. We used a binomial test to statistically quantify the probability that the observed differences in replication fork direction between the various cell lines could be explained by statistical sampling. We applied a threshold of $p < 0.02$ to define a positive result in this test. ($p < 0.02$ was interpreted as evidence of a true difference in directional bias underlying the observed differences).

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cell Culture

An FXS-affected embryo was cultured to the blastocyst stage. HESC line WCMC37 was derived by laser inner cell mass (ICM) dissection of the blastocyst on day 6 as described previously (Thomson et al., 1998). Isolated clumps of ICM cells were plated on PMEF. Outgrowth-containing cells were manually cut and propagated, resulting in a stable culture of undifferentiated hESC. WCMC37 and SI-214 line was fully characterized by stem cell markers *in vitro* and teratoma formation *in vivo*. The use of spare IVF-derived embryos that have been diagnosed as genetically affected for the generation of hESC was approved by Weill Cornell Medical College Institutional review board (Protocol No.0502007737).

H9 (WA09), H14 (WA14), SI-214 and WCMC37 hESCs were grown on PMEF plated at a density of 11,500–13,500 cells/cm2 (GlobalStem, Inc.). ESC were fed daily with hESC media composed of Dulbecco's modified Eagle medium (DMEM)/F12 (11330-032; Gibco), 20% knockout serum replacement (10828-028; Gibco), 3.5 mM glutamine (25030-081; Life Technologies), 0.1 mM MEM non-essential amino acids (NEAA) (11140-050; Gibco), 55 mM 2-mercaptoethanol (21985-023; Life Technologies), and 6 ng/ml fibroblast growth factor 2 (FGF2) (R&D Systems). The last passage before the experiments hESC were grown on Matrigel (BD Biosciences) in conditioned medium. For production of conditioned medium MEF were plated at 50,000 cells/cm^2 in DMEM with 10% FBS overnight. The next day, these media were removed and the cells were washed once with phosphate-buffered saline (PBS). Then hESC media were placed on the cells overnight for conditioning. The next day, the media were removed and FGF2 was added to 10 ng/ml before use. ESC were dissociated with Dispase (1 mg/ml; Worthington) before transfer onto Matrigel. ESC were dissociated into a single-cell suspension with Accutase for 40 min (Innovative Cell Technologies, San Diego, CA) and counted before each experiment.

Human fetal fibroblast GM00011 and GM07072 (Coriell Cell Repositories) were grown in Eagle's Minimum Essential Medium with Earle's salts, NEAA and 15% fetal bovine serum (Invitrogen-Gibco, Carlsbad, CA). Fibroblasts were lifted using Trypsin and passaged every 2-3 days.

Primary human cortex microvascular endothelial cells (MECs; Cell Systems, Rockland, ME; ACBRI 376 lots 2648 and 0537) were originally isolated from the human brain cortex of a healthy 24-year-old women. The cells were elutriated from dispase-dissociated neurons and cultured on gelatin-coated tissue culture dishes in M199 medium supplemented with 20% newborn calf serum (NCS) (both from Invitrogen-Gibco, Carlsbad, CA), 5% human serum (BioCell, Rancho Dominguez, CA), 0.1 g/ml heparin, 0.05 g/ml ascorbic acid (Sigma-Aldrich), 1.6 mM l-glutamine (Invitrogen-Gibco, Carlsbad, CA), Sigma endothelial cell growth factor (Sigma-Aldrich), bovine brain extract (Clonetics BioWhittaker, Walkersville, MD), and 0.05 U/ml penicillin with 0.05 g/ml streptomycin (Invitrogen/Gibco, Carlsbad, CA) to maintain their differentiation state.

Embryoid body (EB) formation

FXS hESC were lifted with Dispase (1 mg/ml; Worthington) and plated in ultra-low

attachment dishes and hESC media without FGF2 for 12 days. Every 2 days the media was changed. FXS EB were dissociated into a single-cell suspension with trypsin for 15 min and counted before each experiments.

Measurement of CGG repeats instability

In order to determine whether CGG repeats are unstable during hESC proliferation, we expanded control (H9) or FXS (SI-214) hESCs from single small colonies and performed southern blotting. To expand hESCs from a single colony, we first cultured hESC as 1000 cells/2 cm^2. ESC were fed with conditioned media that was collected from mouse embryonic fibroblasts (MEF) (MEF CF-1 MITC, GlobalStem # GSC-6001M). MEF were cultured using standard procedures in 20% KSR hESC medium: DMEM:F12, 20% KSR, 4 ng/ml bFGF, 2 mM glutamine, 0.1 mM non-essential amino acids, 50 units/ml penicillin and 50 μg/ml streptomycin, 0.1 mM ß-mercaptoethanol. Conditioned media was collected every 24 hours for 7 day and applied to hESC daily. One colony that contained 6-10 cells from each control and FXS hESC was picked by a needle using a dissection microscope and cultured in MEF-conditioned media on Matrigel for further expansion. The cells expanded from the single colony were passaged further for 3 months (every 7 days) to access CGG repeat expansion by Southern blot analysis.

HESC injection for teratoma formation and teratoma preparation for histology

To evaluate the pluripotent nature of the hESC it is necessary to perform an in-vivo study of teratoma formation. HESC were injected in non-obese diabetic/severe combined immuno-deficient (NOD-SCID) mice that are known to form teratomas rather than reject human cells. To enhance the teratoma formation, approximately 1-5 million hESC cells were injected together with Matrigel (1:1, v/v, total volume 0.5 ml). The WMC-37 hESC and SI-214 ESC were injected subcutaneously in the dorsal back area of the NOD-SCID mice. This area was chosen as less inconvenient because the skin is more mobile and doesn't interfere with the growth of the tumor as much as in testis where tissue growth is constrained by the capsule. After approximately two months, the mouse was sacrificed and the teratoma was processed for histological analysis.

Teratoma tissue was carefully dissected from the host mouse tissue, washed in PBS and fixed overnight in 4% paraformaldehyde (PFA). The next day, the teratoma was extensively washed from the PFA and stored in 30% Ethanol. After this preparation, the teratomas were shipped in Ethanol to the reference histology laboratory (Histoserv, Germantown MA, USA). The standard paraffin embedding and sectioning were performed, followed by haematoxylin and eosin (H&E) staining. Tissue sections were mounted on slides and shipped back for analysis. Histological sections were analyzed using a compound microscope (Nikon, USA) and images were captured with a digital imaging system (Adobe Photoshop, USA). The slides were sent to a pathologist in New York Presbyterian Hospital to confirm the observations.

Flow cytometry

Antibodies against SSEA-3 (conjugated to Alexa 647; BD Biosciences, No. 561145), SSEA-4 (conjugated to Alexa 647; BD Biosciences, No. 560796) and SSEA-1 (Alexa

Fluor® 647; BD Pharmingen No. 560120) were used to quantitate pluripotency markers. In brief, cells were dissociated with Accutase for 40 min, washed with PBS, and then counted before a second centrifugation. One million cells were resuspended in PBS with 0.1% bovine serum albumin (BSA) containing the recommended amount of antibody (for SSEA-3 and for SSEA-4 per test). Cells were incubated in the dark on ice for 1 h before washing 3 times with PBS. Washed cells were resuspended a final time in PBS with 0.1% BSA before analysis on a FACSAria (BD Biosciences).

Immunostaining

For immunostaining cells were grown in 5 well plates. After washing cells with PBS the cells were fixed in 4% paraformaldehyde for 15 min. The cells were permeabilized with PBS/Triton X-100 (0.1%). Primary antibodies used in this study are Oct3/4 (Santa Cruz Biotechnologies, sc-5279; 1:100) and Nanog (BD Pharmingen 560482; 1:200); secondary antibodies conjugated to Alexa488 or 568 (1:1000) were used to visualize primary antibody binding. Cells were incubated for one hour in the dark with the antibodies followed by three wash steps with PBS and incubation with Hoechst 33342 (Molecular Probes H-3570; 1:1,000) to stain nuclei.

FMRP immunostainings were performed by using a peroxidase-based signal amplification system (Tyramid Signal Amplification, TSA, PerkinElmer) according to manufacturer's instructions. TSA enables much higher dilution of primary antibodies than standard protocols. Cells were incubated with antibody against human FMRP (mouse anti-human FMRP, Millipore). Higher dilution of primary antibodies (1:5000 in our experiments) reduced non-specific interactions and improved the specificity of staining. Images were acquired by Zeiss LSM 510 confocal microscopy and processed with a LSM 5 image examiner.

Southern blot

DNA was isolated by Phenol extraction. 12 μg of genomic DNA was digested with 50 U *EagI* and 100 U *EcoRI* enzymes for 4 h at 37°C. Digested genomic DNA was size separated by electrophoresis on a 1 % agarose gel without ethidium bromide (Nolin et al., 2008; Nolin et al., 2003). Enzyme-restricted DNA was blotted onto (Hybond-XL) membrane (Amersham) over night. The membrane was hybridized over night to a PCR-generated probe, using primers; forward, 5′GCTAGCAGGGCTGAAGAGAA, and reverse, 5′CAGTGGAGCTCTCCGAAGTC (PCR product: 595 bp), which was labeled with ^{32}[P]CTP by RadPrime DNA labeling system (Invitrogen). The membrane was washed twice for 5 min with 2 x SSC and once for 8 min at 65°C 2 x SSC/0.5 % SDS. The membrane in Saran Wrap was exposed to Kodak BioMax XAR film at -80°C for 4-7 days.

Single-molecule analysis of replicated DNA (SMARD)

The cells were grown at 37°C for 4 h in the presence of 25 μM 5-iodo-2′-deoxyuridine (Sigma-Aldrich, St. Louis, MO). After washing cells with PBS, ESC medium with 25 μM 5-chloro-2′-deoxyuridine (Sigma-Aldrich, St. Louis, MO) was added to the cultures, and the cells were incubated for an additional 4 h. The cells were lifted with Accutase or Trypsin. Following centrifugation, the cells were resuspended at 3×10^7 cells per ml in

PBS. Melted 1% InCert agarose (Lonza Rockland, Inc., Rockland, ME) in PBS was added to an equal volume of cells at 42°C. The cell suspension was pipetted into a chilled plastic mold with 0.5- by 0.2-cm wells with a depth of 0.9 cm for preparing DNA gel plugs. The gel plugs were allowed to solidify on ice for 30 min. Cells were lysed in buffer containing 1% *n*-lauroylsarcosine (Sigma-Aldrich), 0.5 M EDTA, and 20 mg/ml proteinase K. The gel plugs remained at 50°C for 64 h and were treated with 20 mg/ml proteinase K (Roche Diagnostics), every 24 h. Gel plugs were then rinsed several times with Tris-EDTA (TE) and once with phenylmethanesulfonyl fluoride (Sigma-Aldrich). The plugs were washed with 10 mM $MgCl_2$ and 10 mM Tris-HCl (pH 8.0). The genomic DNA in the gel plugs was digested with 80 units of *SfiI* (New England BioLabs Inc.) at 50°C overnight. The digested gel plugs were rinsed with TE and cast into a 0.7% SeaPlaque GTG agarose gel (Lonza Rockland, Inc.). A gel lambda ladder PFG marker and yeast chromosome PFG marker (both from New England BioLabs, Inc.) were cast next to the gel plugs. A Southern transfer was performed to determine the location of the DNA fragment on the gel. The region of the gel containing the segment of interest was excised and set aside, while the rest of the DNA (which includes the chromosome ladders) was transferred to a membrane (Hybond-XL) and hybridized with a probe located near the CGG repeats (described above). Autoradiography was used to determine the location of the appropriate DNA segment. Gel slices from the appropriate positions in the pulsed-field electrophoresis gel were cut and melted at 72°C for 20 min. GELase enzyme (Epicentre Biotechnologies 1 unit per 50 μl of agarose suspension) was carefully added to digest the agarose and incubated at 45°C for a minimum of 2 h. The resulting DNA solutions were stretched on 3-aminopropyltriethoxysilane (Sigma-Aldrich)-coated glass slides. The DNA was pipetted along one side of a coverslip that had been placed on top of a silane-treated glass slide and allowed to enter by capillary action. The DNA was denatured with sodium hydroxide in ethanol and then fixed with glutaraldehyde.

The slides were hybridized overnight with a biotinylated probe (the blue bars diagrammed on the maps indicate the positions of the probes used). The following day, the slides were rinsed in 2 × SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) 1% SDS and washed in 40% formamide solution containing 2 × SSC at 45°C for 5 min and rinsed in 2 × SSC-0.1% IGEPAL CA-630. Following several detergent rinses (4 times in 4× SSC-0.1% IGEPAL CA-630), the slides were blocked with 1% BSA for at least 20 min and treated with Avidin Alexa Fluor 350 (Invitrogen Molecular Probes) for 20 min. The slides were rinsed with PBS containing 0.03% IGEPAL CA-630, treated with biotinylated anti-avidin D (Vector Laboratories) for 20 min, and rinsed again. The slides were then treated with Avidin Alexa Fluor 350 for 20 min and rinsed again, as in the previous step. The slides were incubated with the IdU antibody, a mouse anti-bromodeoxyuridine (Becton Dickinson Immunocytometry Systems), the antibody specific for CldU, a monoclonal rat anti-bromodeoxyuridine (anti-BrdU) (Accurate Chemical and Scientific Corporation) and biotinylated anti-avidin D for 1 h. This was followed by incubation with Avidin Alexa Fluor 350 and secondary antibodies, Alexa Fluor 568 goat anti-mouse IgG (H+L) (Invitrogen Molecular Probes), and Alexa Fluor 488 goat anti-rat IgG (H+L) (Invitrogen Molecular Probes) for 1 h. After a final PBS/CA-630 rinse, the coverslips were mounted with ProLong gold antifade reagent (Invitrogen).

A Zeiss microscope was used for fluorescence microscopy to follow the nucleoside incorporation of the DNA molecules.

RNA/cDNA Preparation and Quantitative PCR

Cells were lysed in TRIzol reagent (Invitrogen) and RNA was extracted following the TRIzol protocol using manufacturer's instructions. Purified RNA was incubated with DNAse I (RNase-Free DNase Set, Qiagen) on a RNAeasy mini kit column (Qiagen) to digest genomic DNA. cDNA was transcribed using SuperScript III First-Strand Synthesis SuperMix (Invitrogen) in a 20 µl reaction by using1 µg of RNA. Real-Time quantitative PCR reactions were performed using the iQ SYBR Green Supermix (Bio-Rad) and an Eppendorf Mastercycler ep realplex thermocycler. For each reaction 20 ng cDNA was used. Primers used to quantify *FMR1* mRNA and housekeeping gene *GAPDH* mRNA were s follows: *FMR1* F: 5'GTATGGTACCATTTGTTTTTGTG 3', *FMR1* R: 5'CATCATCAGTCACATAGCTTTTTTC 3'; *GAPDH* F: 5' AGCCACATCGCTCAGACACC 3', *GAPDH* R: 5' GTACTCAGCGGCCAGCATCG 3'.

Western blot

For western blotting, hESCs were lysed in RIPA buffer (50 mM Tris [pH 8.0], 150 mM NaCl, 1% Nonidet P-40, 0.1% SDS) containing protease (Roche) and phosphatase (Sigma) inhibitor mixtures. Lysates were clarified by centrifugation for 10 min at 4°C. Proteins were resolved by 4% to 12% SDS/PAGE, transferred to nitrocellulose (Bio-Rad), blocked in 5% (wt/vol) nonfat milk, probed with the mouse anti-FMRP (Millipore) and appropriate secondary antibodies, and detected by enhanced chemiluminescence (GE Life Sciences).

ChIP

ChIP experiments were performed by using EZ-ChIP kit (Millipore, cat: 17-371) and following manufacturer's instructions. For ChIP experiments, approximately 6 million cells were cross-linked with 1% formaldehyde for 15 min. Upon quenching the unreacted formaldehyde with Glycine and extensive washes with PBS, cells were lysed in SDS buffer supplemented with protease inhibitors (supplied in the kit). Cell lysates were sonicated in Bioruptor (Diagenode) with following settings: high power, 30s pulse, 30s interval, 5 min. This was repeated 9 times to get the range of bands between 500-1000 bp for immunoprecipitation or RNA precipitation. Antibodies used in ChIP are: antibody against the methylation at lysine 4 position in Histone 3 (H3K4, Millipore) and antibody against the methylation at lysine 9 position in Histone 3 (H3K9, Millipore). Following the incubation with primary antibodies, antibody/antigen/DNA complex was precipitated by protein G agarose and protein/DNA was eluted by spin columns. Reverse crosslinking was done at 65ºC overnight in the presence of 5 M NaCl. Following the removal of RNA and protein by RNAse A and Proteinase K treatments, respectively, the DNA was used in qRT-PCR with primers that are adjacent to the *FMR1* promoter. DNA primers used in ChIP experiments were as follows: *FMR1* F: GAACAGCGTTGATCACGTGA, *FMR1* R: ACCGGAAGTGAAACCGAAAC; *GAPDH* F: AGGTTTCCAGGAGTGCCTTT, *GAPDH* R: ACCTGATAATTAGGGCAGAC

Fosmids (human genome *GRCh37*/*hg19*)

Fosmid	Start	Ends	from repeats
G248P87323C2	147138275	147177297	downstream
G248P81601D11	147027730	147063703	downstream
G248P87940A10	146973370	147014316	overlaps
G248P80577A2	146917194	146956388	upstream
G248P87545H3	146877388	146917988	upstream
G248P87848F5	146792900	146835948	upstream

Primers (human genome *GRCh37*/*hg19*)

Primer	Position		PCR product
Primer P	146993676	5′GCTAGCAGGGCTGAAGAGAA	595 bp
		5′CAGTGGAGCTCTCCGAAGTC	
Primer 1	146835089	5’TTCAACAGCAGCCAGAATTG	943 bp
		5’GGAGAGCCACATAACCTGGA	
Primer 2	147035438	5’GAGCTGATGCAAG	985 bp
		5’GGCTGCAATCAGTTCAGTCA	

REFERENCES

Nolin, S.L., Ding, X.H., Houck, G.E., Brown, W.T., and Dobkin, C. (2008). Fragile X full mutation alleles composed of few alleles: implications for CGG repeat expansion. Am J Med Genet A *146A*, 60-65.

Nolin, S.L., Dobkin, C., and Brown, W.T. (2003). Molecular analysis of fragile X syndrome. Curr Protoc Hum Genet *Chapter 9*, Unit9 5.

Norio, P., and Schildkraut, C.L. (2001). Visualization of DNA replication on individual Epstein-Barr virus episomes. Science *294*, 2361-2364.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Verlinsky, Y., Strelchenko, N., Kukharenko, V., Rechitsky, S., Verlinsky, O., Galat, V., and Kuliev, A. (2005). Human embryonic stem cell lines with genetic disorders. Reprod Biomed Online *10*, 105-110.